Probability in High Dimensions

Solutions to Exercises, Week 6

August 30, 2020

3.1 Subgaussian variables and Chernoff bounds

Problem 3.1 (Subgaussian variables)**.** *There are several different notions of random variables with a Gaussian tail that are all essentially equivalent up to constants. The aim of this problem is to obtain some insight into these notions.*

a. As a warmup exercise, show that if X is σ^2-subgaussian, then $\operatorname{Var} X \leq \sigma^2$.

b. Show that for any increasing and differentiable function Φ

$$\mathbf{E}\,\Phi(|X|) = \Phi(0) + \int_0^\infty \Phi'(t)\,\mathbf{P}[|X| \geq t]\,dt$$

c. In the following five parts, assume for simplicity that $\mathbf{E}\,X = 0$. Show that if X is σ^2-subgaussian, then $\mathbf{P}[|X| \geq t] \leq 2e^{-t^2/2\sigma^2}$.

d. Show that if $\mathbf{P}[|X| \geq t] \leq 2e^{-t^2/2\sigma^2}$, then $\mathbf{E}[e^{X^2/6\sigma^2}] \leq 2$.

e. Show that if $\mathbf{E}[e^{X^2/6\sigma^2}] \leq 2$, then X is $18\sigma^2$-subgaussian

f. Show that if X is σ^2-subgaussian, then $\mathbf{E}\left[X^{2q}\right] \leq (4\sigma^2)^q q!$ for all $q \in \mathbb{N}$.

g. Show that if $\mathbf{E}\left[X^{2q}\right] \leq (4\sigma^2)^q q!$ for all $q \in \mathbb{N}$, then $\mathbf{E}\left[e^{X^2/8\sigma^2}\right] \leq 2$.

Solution. a. First note that $\psi(0) = \psi'(0) = 0$ for all random variables X, and $\psi''(0) = \operatorname{Var} X$ follows from the expression for $\psi''(\lambda)$ given in the proof of Lemma 3.6. Hence $\psi(\lambda) = \frac{1}{2}\lambda^2 \operatorname{Var} X + o(\lambda^2)$ and so it follows from the definition of subgaussian that $\operatorname{Var} X \leq \sigma^2$.

b. For simplicity assume $X \geq 0$ almost surely. By the Fundamental Theorem of Calculus,

$$\Phi(x) = \Phi(0) + \int_0^x \Phi'(t)\,dt = \Phi(0) + \int_0^\infty \mathbf{I}_{\{t \leq x\}}\,\Phi'(t)\,dt$$

so

$$\begin{aligned}\mathbf{E}\,\Phi(X) &= \Phi(0) + \mathbf{E}\int_0^\infty \mathbf{I}_{\{X \geq t\}}\,\Phi'(t)\,dt \\ &= \Phi(0) + \int_0^\infty \Phi'(t)\,\mathbf{E}\,\mathbf{I}_{\{X \geq t\}}\,dt\end{aligned}$$

by Fubini's/Tonelli's Theorem (the integrand is non-negative by the conditions given), giving the result.

c. Since $\mathbf{E}\,X = 0$, we can apply the Chernoff bound to both the upper and lower tails to obtain

$$\mathbf{P}[X \geq t] \leq e^{-t^2/2\sigma^2} \quad \text{and} \quad \mathbf{P}[X \leq -t] \leq e^{-t^2/2\sigma^2}.$$

For the second bound, we have that $\mathbf{P}[X \leq -t] = \mathbf{P}[-X \geq t]$. Now, we can sum the above inequalities to obtain

$$\mathbf{P}[|X| \geq t] = \mathbf{P}[X \geq t] + \mathbf{P}[-X \geq t] \leq 2e^{-t^2/2\sigma^2}$$

since $t \geq 0$.

d. Take $\Phi(t) = e^{t^2/6\sigma^2}$, which is differentiable and increasing on $\mathbb{R}^+$. Then by b we have

$$\begin{aligned}
\mathbf{E}\,\Phi(|X|) &= 1 + \frac{1}{3\sigma^2}\int_0^\infty te^{t^2/6\sigma^2}\,\mathbf{P}[|X| \geq t]\,dt \\
\mathbf{E}\,e^{X^2/6\sigma^2} &\leq 1 + \frac{1}{3\sigma^2}\int_0^\infty te^{t^2/6\sigma^2}2e^{-t^2/2\sigma^2}\,dt \\
&= 1 + \frac{2}{3\sigma^2}\int_0^\infty te^{-t^2/3\sigma^2}\,dt \\
&= 1 + 1 = 2.
\end{aligned}$$

e. Recall that we assume $\mathbf{E}[X] = 0$. We have for any λ,

$$\begin{aligned}
\mathbf{E}[e^{\lambda X}] &= \mathbf{E}[\sum_{k\geq 0}\frac{(\lambda X)^k}{k!}] \\
&= 1 + \frac{\lambda^2}{2}\,\mathbf{E}[X^2\sum_{k\geq 2}2\frac{(\lambda X)^{k-2}}{k!}] \\
&\leq 1 + \frac{\lambda^2}{2}\,\mathbf{E}[X^2\sum_{k\geq 0}\frac{(|\lambda X|)^k}{k!}\frac{2}{(k+1)(k+2)}] \\
&\leq 1 + \frac{\lambda^2}{2}\,\mathbf{E}[X^2\sum_{k\geq 0}\frac{(|\lambda X|)^k}{k!}] \\
&= 1 + \frac{\lambda^2}{2}\,\mathbf{E}[X^2e^{|\lambda X|}]
\end{aligned}$$

Let $c > 0$. Using Young's inequality we have,

$$|\lambda X| \leq \frac{c\lambda^2}{2} + \frac{X^2}{2c}$$

Therefore,

$$\begin{aligned}
\mathbf{E}[e^{\lambda X}] &\leq 1 + \frac{\lambda^2}{2}e^{\frac{c\lambda^2}{2}}\,\mathbf{E}[X^2e^{\frac{X^2}{2c}}] \\
&\leq 1 + \frac{c\lambda^2}{2}e^{\frac{c\lambda^2}{2}}\,\mathbf{E}[e^{\frac{X^2}{c}}] \quad \text{(using the fact that } 2\frac{X^2}{2c} \leq e^{\frac{X^2}{2c}}) \\
&\leq 1 + c\lambda^2e^{\frac{c\lambda^2}{2}} \quad \text{(by choosing } c = 6\sigma^2)
\end{aligned}$$

Now we use the following Lemma.

Lemma 1. For $x \geq 0$, we have :

$$1 + 2xe^x \leq e^{2x}$$

Proof. It is sufficient to show the inequality between the coefficients in the series expansion of both sides. For the left-hand side, the coefficient of x^k is $\frac{2}{(k-1)!}$, and for the right-hand expression, the coefficient of x^k is $\frac{2^k}{k!}$. So the inequality is equivalent to having for any $k \geq 1$:

$$\begin{aligned} \frac{2}{(k-1)!} &\leq \frac{2^k}{k!} \\ \iff 2k &\leq 2^k \\ \iff k &\leq 2^{k-1} \end{aligned}$$

which is true for any $k \geq 1$. □

Using this we now have,

$$\begin{aligned} \mathbf{E}[e^{\lambda X}] &\leq 1 + c\lambda^2 e^{\frac{c\lambda^2}{2}} \\ &\leq e^{\frac{2c\lambda^2}{2}} \\ &= e^{\frac{12\sigma^2\lambda^2}{2}} \end{aligned}$$

We conclude that X is $12\sigma^2$-subgaussian and thus $18\sigma^2$-subgaussian.

f. Let $\Phi(x) = x^{2q}$ for $q \in \mathbb{N}$. Then $\Phi'(x) = 2qx^{2q-1}$. Since Φ is differentiable and increasing, we will use part (b), i.e.

$$\begin{aligned} \mathbf{E}[X^{2q}] &= \Phi(0) + \int_0^\infty 2qt^{2q-1}\, \mathbf{P}(|X| \geq t) dt \\ &\leq 4q \int_0^\infty t^{2q-1} e^{-t^2/2\sigma^2} dt. \end{aligned}$$

Now, performing the change of variable $u = \frac{t^2}{2\sigma^2}$, we can rewrite the inequality as

$$\begin{aligned} \mathbf{E}[X^{2q}] &\leq 4q \int_0^\infty (2\sigma^2 u)^{q-1} e^{-u} \sigma^2 du \\ &= 2^{q+1}\sigma^{2q} \int_0^\infty u^{q-1} e^{-u} du \\ &= 2^{q+1}\sigma^{2q} q! \\ &\leq (4\sigma^2)^q q!, \end{aligned}$$

where the second last line holds from the definition of the gamma function evaluated at a natural number.

g. We can prove this directly by expanding $\mathbf{E}\left[e^{X^2/8\sigma^2}\right]$ as a power series:

$$\begin{aligned}
\mathbf{E}\left[e^{X^2/8\sigma^2}\right] &= \sum_{q=0}^{\infty} \frac{\mathbf{E}[X^{2q}]}{8^q \sigma^{2q} q!} \\
&\leq \sum_{q=0}^{\infty} \frac{(4\sigma^2)^q q!}{8^q \sigma^{2q} q!} \\
&= \sum_{q=0}^{\infty} 2^{-q} = 2.
\end{aligned}$$

Problem 3.2 (Tightness of Hoeffding's Lemma)**.** *Show that the bound of Hoeffding's lemma is the best possible by considering $P(X = a) = P(X = b) = \frac{1}{2}$.*

Solution. The distribution suggested has variance $(b - a)^2/4$, so the tightness follows from Problem 3.1a.

Problem 3.3 (Chernoff bound vs. moments)**.** *Show that for $t \geq 0$,*

$$\mathbf{P}[X - \mathbf{E}\,X \geq t] \leq \inf_{p \geq 0} \frac{\mathbf{E}\left[(X - \mathbf{E}\,X)_+^p\right]}{t^p} \leq \inf_{\lambda \geq 0} e^{-\lambda t}\, \mathbf{E}\left[e^{\lambda(X - \mathbf{E}\,X)}\right].$$

Solution. We can see that, for any $\lambda \geq 0$,

$$\begin{aligned}
\mathbf{E}\left[e^{\lambda(X - \mathbf{E}\,X)}\right] &\geq \mathbf{E}\left[\mathbf{1}_{X - \mathbf{E}\,X > 0}\; e^{\lambda(X - \mathbf{E}\,X)}\right] \\
&= \mathbf{E}\left[\mathbf{1}_{X - \mathbf{E}\,X > 0} \sum_{p=0}^{\infty} \frac{\lambda^p (X - \mathbf{E}\,X)^p}{p!}\right] \\
&= \mathbf{E}\left[\sum_{p=0}^{\infty} \frac{\lambda^p (X - \mathbf{E}\,X)_+^p}{p!}\right] \\
&= \sum_{p=0}^{\infty} \frac{\lambda^p\, \mathbf{E}\left[(X - \mathbf{E}\,X)_+^p\right]}{p!},
\end{aligned}$$

where the final line follows since the terms $(X - \mathbf{E}\,X)_+^p$ are non-negative for all $p \geq 0$. Therefore, we have

$$\begin{aligned}
e^{-\lambda t}\, \mathbf{E}\left[e^{\lambda(X - \mathbf{E}\,X)}\right] &\geq \left(\sum_{p=0}^{\infty} \frac{\lambda^p\, \mathbf{E}\left[(X - \mathbf{E}\,X)_+^p\right]}{p!}\right) \Bigg/ \left(\sum_{p=0}^{\infty} \frac{\lambda^p t^p}{p!}\right) \\
&\geq \inf_{p \geq 0} \left\{ \frac{\lambda^p\, \mathbf{E}\left[(X - \mathbf{E}\,X)_+^p\right]}{p!} \Bigg/ \frac{\lambda^p t^p}{p!} \right\} \\
&= \inf_{p \geq 0} \frac{\mathbf{E}\left[(X - \mathbf{E}\,X)_+^p\right]}{t^p},
\end{aligned}$$

where the second inequality follows since the terms in both sums are all non-negative and the sums converge. Since the right-hand side of our final result is independent of λ, we could optimize the left-hand side with respect to λ to obtain the final result.

Solution (alternative). Starts similarly to the previous one, but the end is different. For any $\lambda \geq 0$,

$$\begin{aligned}
\mathbf{E}[e^{\lambda(X-\mathbf{E}[X])}] &\geq \mathbf{E}\left[\mathbf{1}_{X-\mathbf{E}[X]>0}\ e^{\lambda(X-\mathbf{E}[X])}\right] \\
&= \mathbf{E}\left[\mathbf{1}_{X-\mathbf{E}[X]>0}\ \sum_{p=0}^{\infty}\frac{(\lambda(X-\mathbf{E}[X]))^p}{p!}\right] \\
&= \sum_{p=0}^{\infty}\frac{\lambda^p}{p!}\,\mathbf{E}\left[\mathbf{1}_{X-\mathbf{E}[X]>0}\ (X-\mathbf{E}[X])^p\right] \\
&= \sum_{p=0}^{\infty}\frac{(\lambda t)^p}{p!}\,\frac{\mathbf{E}\left[(X-\mathbf{E}[X])_+^p\right]}{t^p} \\
&\geq \sum_{p=0}^{\infty}\frac{(\lambda t)^p}{p!}\,\inf_{p\geq 0}\frac{\mathbf{E}\left[(X-\mathbf{E}[X])_+^p\right]}{t^p}.
\end{aligned}$$

Therefore,

$$e^{-\lambda t}\,\mathbf{E}[e^{\lambda(X-\mathbf{E}[X])}] \geq \inf_{p\geq 0}\frac{\mathbf{E}\left[(X-\mathbf{E}[X])_+^p\right]}{t^p},$$

which yields the result given that the inequality holds for all $\lambda \geq 0$.
Remark: We should consider $t > 0$.

Problem 3.4.

Problem 3.5 (Maxima of subgaussian variables)**.** *Let $X_1, X_2, \ldots$ be (not necessarily independent) σ^2-subgaussian random variables. Show that*

$$\mathbf{P}\left[\max_{i\leq n}\{X_i - \mathbf{E}[X_i]\} \geq (1+\epsilon)\sigma\sqrt{2\log n}\right] \xrightarrow{n\to\infty} 0 \quad \textit{for all } \epsilon > 0.$$

Solution. We have

$$\begin{aligned}
\mathbf{P}\left[\max_{i\leq n}\{X_i - \mathbf{E}[X_i]\} \geq (1+\epsilon)\sigma\sqrt{2\log n}\right] &= \mathbf{P}\left[\bigcup_{i=1}^{n}\left\{X_i - \mathbf{E}[X_i] \geq (1+\epsilon)\sigma\sqrt{2\log n}\right\}\right] \\
&\leq \sum_{i=1}^{n}\mathbf{P}\left[X_i - \mathbf{E}[X_i] \geq (1+\epsilon)\sigma\sqrt{2\log n}\right] \\
&\leq \sum_{i=1}^{n} 2\exp\left\{-\left((1+\epsilon)\sigma\sqrt{2\log n}\right)^2 \Big/ (2\sigma^2)\right\} \\
&= 2n\frac{1}{n^{(1+\epsilon)^2}} \xrightarrow{n\to\infty} 0,
\end{aligned}$$

for all $\epsilon > 0$.